

## **Using the IRAP to Explore Natural Language Statements**

Deirdre Kavanagh, Ian Hussey, Ciara McEnteggart, Yvonne Barnes-Holmes, and Dermot  
Barnes-Holmes

α

Experimental-Clinical and Health Psychology, Ghent University, Ghent, Belgium,

[deirdrekavanagh@ugent.be](mailto:deirdrekavanagh@ugent.be)

[ian.hussey@ugent.be](mailto:ian.hussey@ugent.be)

[ciara.mcenteggart@ugent.be](mailto:ciara.mcenteggart@ugent.be)

[yvonne.barnes-holmes@ugent.be](mailto:yvonne.barnes-holmes@ugent.be)

[dermot.barnes-holmes@ugent.be](mailto:dermot.barnes-holmes@ugent.be)

### **Author Note**

Correspondence should be addressed to Deirdre Kavanagh, Experimental-Clinical and Health  
Psychology, Ghent University, Ghent, Belgium, [deirdrekavanagh@ugent.be](mailto:deirdrekavanagh@ugent.be)

### **Abstract**

This study explored a modification to the typical presentation of label and target stimuli on Implicit Relational Assessment Procedure (IRAP) effects. We asked whether combining the labels and targets into a single phrase would influence performances. The key purpose of the study was to determine the feasibility of altering the way in which stimuli are presented within the IRAP, so as to potentially employ more complex natural language-like statements in future research. In the Typical IRAP employed here, labels and targets were presented as separate words, while in the Natural Language IRAP they were combined to form a single statement. The results demonstrated no substantive differences in the effects recorded on both types of IRAP, thus supporting the future use of a Natural Language version.

*Keywords:* RFT; Implicit Relational Assessment Procedure; Natural Language IRAP

The purpose of this brief empirical report is to demonstrate the feasibility and potential of a novel variant of the Implicit Relational Assessment Procedure (IRAP) that allows the researcher to employ natural language statements as stimuli. The IRAP emerged directly from Relational Frame Theory (RFT) as a methodology to assess verbal relations (Barnes-Holmes, Hayden, Barnes-Holmes, & Stewart, 2008; Hussey, Barnes-Holmes, & Barnes-Holmes, 2015). The procedure has shown utility in the study of many forms of psychological suffering (see Vahey, Nicholson, & Barnes-Holmes, 2015, for a recent meta-analysis). Traditionally, an IRAP pairs label stimuli (e.g., “beetles”) presented at the top of the screen with target stimuli (e.g., “delicious”) presented in the middle. These label and stimulus pairs form trial-types that are generally analyzed separately (e.g., the difference in response latencies between responding “True” versus “False” to the stimulus pairing beetles-delicious).

Several recent studies have employed full statements as stimuli, rather than single words or pictures within the traditional IRAP (e.g., Hussey & Barnes-Holmes, 2012; Nicholson & Barnes-Holmes, 2012; Remue, De Houwer, Barnes-Holmes, Vanderhasselt, & De Raedt, 2013). Indeed, one of the core reasons for using the IRAP is that it can readily accommodate full statements (Gawronski & De Houwer, 2014; see also De Houwer, Heider, Spruyt, Roets, & Hughes, 2015). While it is possible to use statements within the IRAP as it stands, these must be divided into label stimuli (e.g., “I want to be”) and target stimuli (e.g., “valuable”, see Remue et al.). As such, the presentation of these “divided” statements may appear awkward or unconventional, relative to statements written in natural language. The purpose of this brief report is to highlight preliminary data that we have gathered using a version of the IRAP that involves presenting single whole statements in a natural language format. We believe that publishing the results of the current study is particularly timely because at the time of writing our

## NATURAL LANGUAGE IRAP

research group were close to releasing a new and considerably up-graded version of the IRAP program that provides the user with an option to present label (e.g., “beetles”) and target (e.g., “tasty”) stimuli in a natural language format (e.g., “beetles are tasty”).

One way in which the Natural Language IRAP might be of benefit to researchers would be in attempting to insert questionnaire-based items into the procedure. When this general approach is adopted, the fact that the presentation of targets and labels is randomized across trials in a typical IRAP may present a problem. Imagine, for example, that a researcher wanted to assess responses to two statements, “Social events make me feel anxious” and “Criticism makes me feel depressed”. In the typical IRAP, two labels might be used, “social events” and “criticism”, along with two targets, “makes me feel anxious” and “makes me feel depressed”. In running the IRAP, however, the designated label and target stimuli may not necessarily occur together. For example, across trials both the “anxious” and “depressed” targets would appear with both of the labels. As such, the IRAP would be assessing the extent to which social events and criticism make you feel both depressed and anxious. Obviously there may be theoretical or conceptual reasons for avoiding conflating anxiety and depression in this way. A Natural Language version of the IRAP in which a single statement is inserted would circumvent this problem. A similar approach has been adopted with the development of the Relational Response Task (RRT, De Houwer, et al., 2015).

In the current study, we exposed participants to a traditional IRAP and to a natural-language version of the same program, counterbalancing the order of presentation. The main focus of our analyses was to determine if any substantive differences in the individual *D*-IRAP scores, attrition rates, and length of time to complete the procedures would be observed between

the two types of IRAP. Given that this was the first study of its kind, and was therefore largely exploratory, we refrained from making any specific predictions.

### **Method**

#### **Participants**

Twenty-two undergraduates (15 female, 9 male) were recruited from the National University of Ireland Maynooth (NUIM). Ages ranged from 19 to 28 years ( $M = 20.2$ ,  $SD = 2.2$ ).

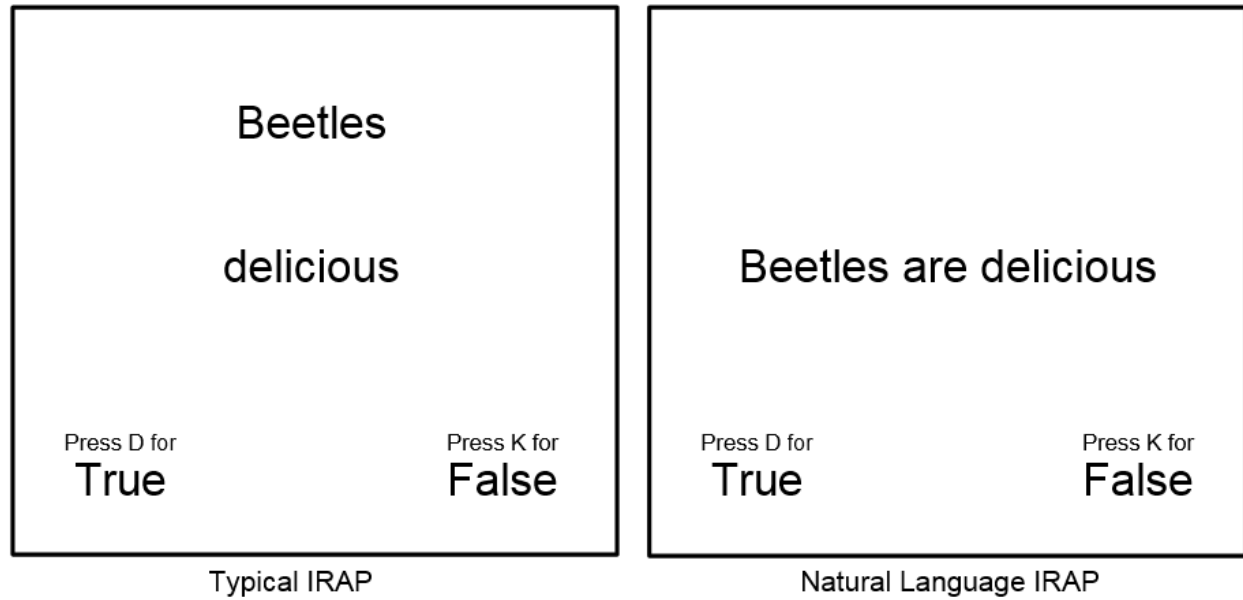
#### **Setting**

All aspects of the research were conducted in a laboratory at NUIM. All participation was individual. The researcher was in the laboratory during instructional and practice phases of the IRAP. Participation lasted 30 minutes, with scheduled breaks if needed.

#### **Apparatus and Materials**

All aspects of the experiment were automated. The study involved two IRAPs -- the *Typical IRAP* and the *Natural Language IRAP*.

**The Typical IRAP.** The Typical IRAP was referred to as such because its screen format was identical to almost all published IRAPs (see Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010). That is, the label stimulus appeared in the top center of the screen, with the target below, and two static response options on the bottom left- and right-hand sides, see Figure 1 (left-hand side).



*Figure 1. A comparison of the Typical IRAP and the Natural Language IRAP trials.*

The Typical IRAP presented 12 labels. Six were fruits (e.g., “apricots”) and six were insects (e.g., “centipedes”), adapted from Nosek and Banaji (2001, see Table 1). This IRAP also presented 12 targets. Six were positive words (e.g., “sweet”) and six were negative (e.g., “rotten”). Two static response options (“True” and “False”) were presented at the bottom left- and right-hand corners, respectively.

Table 1

*Stimuli employed as the Typical IRAP’s labels and targets.*

Label stimuli		Target stimuli	
Fruits	Insects	Positive	Negative
Apricots	Centipedes	Juicy	Rotten
Peaches	Cockroaches	Sweet	Nasty
Raspberries	Maggots	Appetizing	Terrible
Watermelon	Spiders	Tasty	Revolting
Grapes	Wasps	Delicious	Foul
Blueberries	Beetles	Enjoyable	Disgusting

**The Natural Language IRAP.** The Natural Language IRAP was referred to as such because its screen format differed from a typical IRAP in that the label and target stimuli were combined to form a sentence or statement as it would typically appear in natural language. Thus in the current case, on each trial, the label and target stimuli were combined with “are” to form a short statement in the center of the screen, see Figure 1 (right-hand side). Consider the Typical IRAP in which the label “beetles” appeared above the target “delicious”. In contrast, in the Natural Language IRAP, these two stimuli were combined with “are” to form the statement “beetles are delicious”. All of the stimuli used in the Natural Language IRAP were identical to the Typical IRAP (see Table 2). The valence of the stimuli was not formally tested, but was similar to that employed in previously published studies of the Go/No-go Association Task (Nosek & Banaji, 2001).

Table 2

*Stimuli employed as the Natural Language IRAP’s statements for each trial-type.*

Trial-type	Statement
Fruits-Positive	Apricots are juicy.
	Blueberries are sweet.
	Peaches are appetizing.
	Raspberries are tasty.
	Watermelon is delicious.
Fruits-Negative	Grapes are enjoyable.
	Apricots are nasty.
	Blueberries are foul.
	Peaches are revolting.
	Raspberries are disgusting.
Insects-Positive	Watermelon is rotten.
	Grapes are terrible.
	Centipedes are tasty.
	Cockroaches are sweet.
	Maggots are enjoyable.
	Spiders are appetizing.

Wasps are juicy. Beetles are delicious.	
Insects-Negative	Centipedes are rotten. Cockroaches are nasty. Maggots are terrible. Spiders are revolting. Wasps are foul. Beetles are disgusting.

## Procedure

All procedures in the current study were in accordance with the ethical standards of the institutional research committee, and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants. The experimental sequence comprised two IRAPs, the order of which was counterbalanced across participants. The section below describes the procedure for participants exposed to the Typical IRAP first and the Natural Language IRAP thereafter. The length of time taken by participants to complete each of the IRAPs was also recorded.

**The Typical IRAP.** Prior to the first practice block, participants were verbally instructed that each trial would present a word on top, with a word in the center, and that their task was to respond with “True” or “False” in accordance with the rule presented at the beginning of the block (see below). Participants were informed that the rule would switch during the next block, so that they would then respond in the opposite manner. These instructions also highlighted the criterion for accurate (i.e., >80%) and fast (i.e., <2,000 ms) responding.

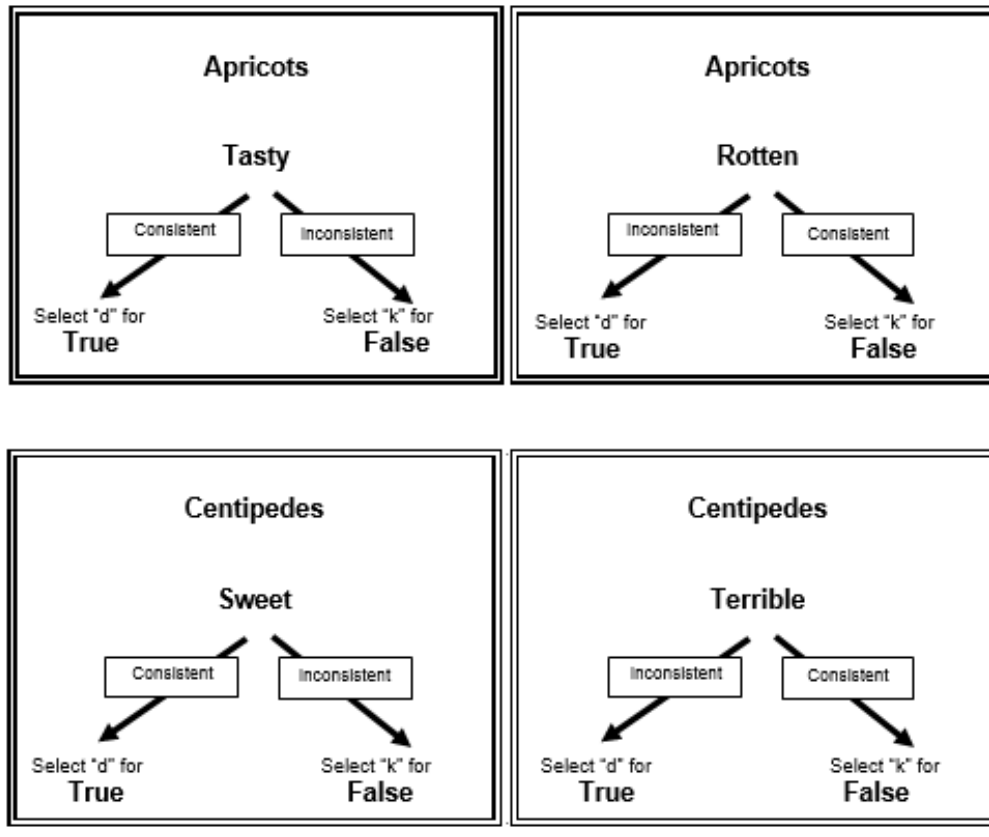
The IRAP consisted of blocks of 24 trials presented quasi-randomly. On each trial, a label (e.g., “apricots”) appeared at the top, a target (e.g., “tasty”) in the middle, and the response options (“True” and “False”) on the bottom left- and right-hand corners (see Figure 1, left panel). Participants selected a response by pressing “D” (for the left option) or “K” (for the right). If a



participant chose the correct response, the screen cleared, and the next trial appeared. If the participant chose incorrectly, a red “X” appeared until a correct response was emitted.

The feedback contingencies for IRAP blocks alternated according to the rule specified at the beginning of each block. The Typical IRAP comprised two rules for responding. One rule was consistent with likely existing verbal relations (“Fruits taste good and insects taste bad”), while the other rule was inconsistent with these (“Fruits taste bad and insects taste good”). Hence, correct responding involved switching between rules from block to block. The order in which the two types of blocks were presented was counterbalanced across participants.

The Typical IRAP comprised four trial-types: Fruits-Positive; Fruits-Negative; Insects-Positive; and Insects-Negative (see Figure 2). During blocks of trials in which the rule was consistent with existing verbal relations, the following responses were deemed correct: Fruits-Positive/True; Fruits-Negative/False; Insects-Positive/False; Insects-Negative/True. During blocks of trials in which the rule was *inconsistent* with existing verbal relations, the following responses were deemed correct: Fruits-Positive/False; Fruits-Negative/True; Insects-Positive/True; Insects-Negative/False.



*Figure 2.* Examples of the four trial-types in the Typical IRAP. On each trial, a label stimulus (Fruits or Insects), a target stimulus (Positive or Negative), and two response options (“True” and “False”) appeared on-screen simultaneously. This generated four trial-types: Fruits-Positive (True); Fruits-Negative (False); Insects-Positive (False); and Insects-Negative (True). The words ‘Consistent’ and ‘Inconsistent’ were not shown on-screen

The IRAP commenced with a minimum of one pair of practice blocks. If participants failed to achieve both accuracy and latency criteria across a pair of blocks, they received automated feedback, and practice blocks continued to a maximum of four pairs of blocks. Failing to meet the criteria after four pairs of practice blocks terminated participation and these data were discarded. When the criteria were reached on a pair of practice blocks, participants proceeded automatically to three pairs of test blocks. No performance criteria were employed for participants to progress across the three pairs of test blocks, but performance feedback was presented at the end of each block to encourage participants to maintain the criteria. The program

## NATURAL LANGUAGE IRAP

automatically recorded response accuracy (based on the first response emitted on each trial) and response latency (time in ms between trial onset and emission of correct response) on each trial.

**The Natural Language IRAP.** Participants were verbally instructed that each trial would present a single statement in the center of the screen and that their task was to respond with “True” or “False” in accordance with the rule presented at the beginning of the block. All other instructions and parameters of the Natural Language IRAP were identical to those outlined above for the Typical IRAP.

## Results

### IRAP Data

All aspects of data processing for the IRAP adhered to standard conventions (e.g., Nicholson & Barnes-Holmes, 2012). One participant failed to meet the mastery criteria on the Natural Language IRAP practice blocks and was therefore excluded from the analysis. It was intended that the data from participants who failed to maintain the mastery criteria across two test blocks would be excluded from analysis. However, no data were excluded on this basis. Therefore, the overall number of participants that met the pass criteria for both IRAPs was 21. In addition, the *D*-scores from the Insects-Positive and Insects-Negative trial-types were inverted (i.e., multiplied by -1) to create a common axis of comparison across the four trial-types (see Hussey, Thompson, McEnteggart, Barnes-Holmes, and Barnes-Holmes, 2015). As a result, positive *D*-scores indicated responding True more quickly than False when presented with Fruits-Positive and Insects-Positive, and responding False more quickly than True when presented with Fruits-Negative and Insects-Negative. Negative *D*-scores were indicative of the opposite pattern (e.g., responding False more quickly than True when presented with Fruits-Positive). In effect,

positive *D*-scores indicated a positive bias to fruits and/or insects, whereas negative scores indicated a negative bias.

The mean *D*-scores for each trial-type for each IRAP are illustrated in Figure 3. Both IRAPs produced similar effects across the four trial-types. Three of the trial-types produced positive biases, with the strongest observed for Fruits-Positive. The effects for the Insects-Negative trial-type were negligible. The *D*-scores were subjected to a repeated measures 4×2 ANOVA. There was a main effect for trial-type [ $F(3,60) = 14.361, p < .001, \eta p^2 = .42$ ], but the effects for IRAP type and the interaction were both non-significant (both  $ps > .58$ ). Four paired *t*-tests confirmed that none of the four trial-type *D*-scores differed significantly between the two IRAPs (all  $ps > .32$ ). The absence of any difference is unlikely due to insufficient power, given that a recent meta-analysis of IRAP effects indicated that only 8-10 participants are required to achieve power of 0.8 when using repeated measures *t*-tests (Vahey et al., 2015).

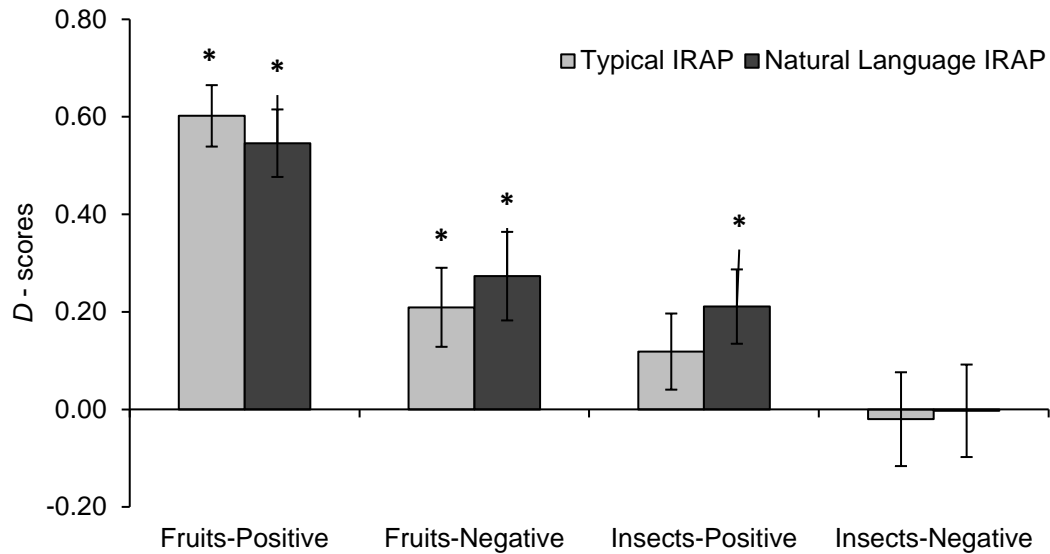


Figure 3. Mean *D*-scores on the Typical and Natural Language IRAP trial-types. Positive *D*-scores indicate positive bias and negative *D*-scores indicate negative bias. \* indicates *D*-scores which are significantly different from zero.

Post-hoc tests, collapsing across IRAP types, indicated that all comparisons except one (Fruits-Negative vs. Insects-Positive,  $p > .4$ ), were significant or marginally so (all other  $ps < .06$ ). One sample *t*-tests indicated that responding on Fruits-Positive was significant on both IRAPs (Typical IRAP:  $M = .60$ ,  $SD = .29$ ,  $t(20)=9.55$ ,  $p < .001$ ; Natural Language IRAP:  $M = .55$ ,  $SD = .32$ ,  $t(20)=7.87$ ,  $p < .001$ ), as was responding on Fruits-Negative (Typical IRAP:  $M = .21$ ,  $SD = .37$ ,  $t(20)=2.57$ ,  $p = .02$ ; Natural Language IRAP:  $M = .27$ ,  $SD = .42$ ,  $t(3.02)=$ ,  $p = .01$ ). Responding on Insects-Positive was also significant on the Natural Language IRAP ( $M = .21$ ,  $SD = .35$ ,  $t(20)=-2.76$ ,  $p = .01$ ; all other  $ps > .15$ ).

Finally, a dependent measures *t*-test was used to determine if the two IRAPs differed in terms of the time taken to complete (Typical IRAP:  $M = 9.84$  min,  $SD = 3.21$ ; Natural Language IRAP:  $M = 9.50$  min,  $SD = 2.50$ ), but this test proved to be non-significant ( $p = .57$ ). As noted above, only one participant out of 22 who started the experiment failed to complete both IRAPs,

and thus it was not possible to make a meaningful comparison of attrition rates across the two procedures.

### **Discussion**

The purpose of the current study was to demonstrate the feasibility of altering the way in which stimuli are presented within the IRAP so as to employ natural language-like statements. The main strategy adopted here was to compare performance on a Natural Language versus a Typical IRAP to determine if any substantive differences would emerge. The data indicated that no significant differences emerged between the two procedures, thus suggesting that a Natural Language IRAP could be used for research in which relatively complex verbal stimuli need to be presented. On balance, it must be noted that this is a preliminary study that focused on the *D*-scores per se, and thus further research is needed to determine if the two procedures differ in terms of predictive validity.

In reflecting upon the two procedures, it is worth noting again a subtle but important difference between the two IRAPs. Specifically, the Typical IRAP involves presenting separate labels and targets that are quasi-randomly mixed. In the current study, for example, the label “Apricots” could, in principle, appear with any of the 12 target words. In contrast, the Natural Language IRAP involved presenting 24 statements in which the same label and target stimuli always appeared together (e.g., “apricots are juicy”). As noted above, no significant differences emerged between the two IRAPs, and thus this procedural difference did not appear to impact substantively on the observed performances. However, this difference may be important in other domains, such as that outlined in the anxiety/depression example provided in the Introduction. On balance, if a researcher wanted to maintain the mixing of labels and targets in a natural language format, it may be useful in future studies employing a Natural Language IRAP to

generate a large pool of statements from which the program selects quasi-randomly, ensuring that the label and target stimuli can appear in any combination.

It is important to note that failing to find a statistically significant difference between the two IRAPs does not indicate that they are functionally equivalent. Indeed, there may be contexts in which a Natural Language IRAP encourages participants to respond to an entire label or target, in a way that a Typical IRAP does not. A possible example (Drake, Timko, & Luoma, 2016) is provided by a recently published study that presented just two labels (“I am willing to have” and “I try to get rid of”) with multiple targets (anxiety-relevant emotions, “anxiety”, “fear”, “worry” and positive emotions, “contentment”, “happy”, and “relaxation”). Given that participants were required to respond in under 2,000ms. (as is standard practice in IRAP studies), it is possible or perhaps even likely that at least some participants responded only to the first two words of each label to discriminate successfully between them. If this occurred, participants would have read the trial-type “I am willing to have anxiety”, for example, as “I am anxiety”. As such, responding “True” to this trial-type would render such a response consistent with fusion, rather than *defusion*. Interestingly, this was the nature of the correlation that emerged from the study (i.e., responding “True” more quickly than “False” predicted lower defusion and lower acceptance on explicit measures). In other words, confirming rather than denying “I am anxiety/fear/worry” predicted lower scores on the Drexel Defusion Scale and The Acceptance and Action Questionnaire. Of course, this interpretation of the results of the Drake et al. study remains highly speculative at the current time, but it would be interesting to repeat the study using a Natural Language IRAP which may discourage participants from responding to only part of a label that is presented separately at the top half of the screen.

In closing, two points are worth noting. First, the attrition rates were low for both IRAPs in the current study (cf. Hughes & Barnes-Holmes, 2012, Table 1). One possible reason for the lack of attrition may be the relative simplicity of the stimuli that were employed and the detailed rules presented before each block. Furthermore, it is possible that some of the participants may have been involved in a previous IRAP study and thus the full benefit of a natural language format may not have emerged in terms of reducing attrition rates with completely IRAP-naïve participants.

Second, the size of the IRAP effects are relatively uneven across the four trial-types. Perhaps most strikingly, the effect for Fruits-Positive was exceptionally strong, whereas the effect for Insects-Negative was virtually absent. This could be interpreted as indicating that participants had very positive attitudes toward fruits, but were ambivalent towards insects. Intuitively, this seems like an odd result. On balance, recent research from our group has highlighted that such unusual effects may be attributable, at least in part, to the provision of very specific rules presented at the beginning of each block of trials (Finn, Barnes-Holmes, Hussey, & Graddy, 2016). While this remains an interesting avenue that we and other researchers will likely pursue, it remains the case that both IRAPs in the current study were similarly affected by this variable.



### **Acknowledgements**

The data for the current manuscript was collected at the National University of Ireland, Maynooth, and was prepared with the support of the FWO Type I Odysseus Programme at Ghent University, Belgium.

### References

- Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. *The Psychological Record*, 60, 527–542.
- Barnes-Holmes, D., Hayden, E., Barnes-Holmes, Y. & Stewart, I. (2008). The Implicit Relational Assessment Procedure (IRAP) as a response-time and event-related-potentials methodology for testing natural verbal relations: A preliminary study, *The Psychological Record*, 58, 497-516.
- De Houwer, J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). The relational responding task: Toward a new implicit measure of beliefs. *Frontiers in Psychology*, 6, 1-9.
- Drake, C. E., Timko, C. A., & Luoma, J. B. (2016). Exploring an implicit measure of acceptance and experiential avoidance of anxiety. *The Psychological Record*, 66, 463-475.
- Finn, M., Barnes-Holmes, D., Hussey, I., & Graddy, J. (2016). Exploring the behavioural dynamics of the implicit relational assessment procedure: The impact of three types of introductory rules. *The Psychological Record*, 60, 309-321.
- Gawronski, B., & De Houwer, J. (2014). Implicit measures in social and personality psychology. In H. T. Reis, & C. M. Judd (2<sup>nd</sup> Ed), *Handbook of research methods in social and personality psychology* (pp. 283-310). New York: Cambridge University Press.
- Hughes, S., & Barnes-Holmes, D. (2012). A functional approach to the study of implicit cognition: The Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. In S. Dymond & B. Roche (Eds.), *Advances in Relational Frame Theory: Research and application* (pp. 97–125). Oakland CA: New Harbinger.

- Hussey, I., & Barnes-Holmes, D. (2012). The Implicit Relational Assessment Procedure as a measure of implicit depression and the role of psychological flexibility. *Cognitive and Behavioral Practice*, 19, 573-582.
- Hussey, I., Barnes-Holmes, D., & Barnes-Holmes, Y. (2015). From Relational Frame Theory to implicit attitudes and back again: Clarifying the link between RFT and IRAP research. *Current Opinion in Psychology*, 2, 11-15.
- Hussey, I., Thompson, M., McEnteggart, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2015). Interpreting and inverting with less cursing: A guide to interpreting IRAP data. *Journal of Contextual Behavioral Science*, 4, 157-162.
- Nicholson, E., & Barnes-Holmes, D. (2012). The Implicit Relational Assessment Procedure (IRAP) as a measure of spider fear. *The Psychological Record*, 62, 263–278.
- Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association task. *Social Cognition*, 19, 625-664.
- Remue, J., De Houwer, J., Barnes-Holmes, D., Vanderhasselt, M. A., & De Raedt, R. (2013). Self-esteem revisited: Performance on the implicit relational assessment procedure as a measure of self- versus ideal self-related cognitions in dysphoria. *Cognition & Emotion*, 27, 1441-1449.
- Vahey, N. A., Nicholson, E., & Barnes-Holmes, D. (2015). A meta-analysis of criterion effects for the Implicit Relational Assessment Procedure (IRAP) in the clinical domain. *Journal of Behavior Therapy and Experimental Psychiatry*, 48, 59–65.
- <http://doi.org/10.1016/j.jbtep.2015.01.004>

